POW 2020-16

2016____ Chae Jiseok

September 29, 2020

We use * to denote conjugate transpose. Let two integers n and k be given so that $1 \leqslant k \leqslant n$. For an $n \times n$ Hermitian matrix A let f denote the function given in the problem, that is, let

$$f(A) = \lambda_1(A) + \cdots + \lambda_k(A).$$

We make a few claims.

Claim 1. *For an $n \times n$ Hermitian matrix A, let $u_1, \cdots, u_n$ be unit norm eigenvectors correspoding to $\lambda_1(A), \cdots, \lambda_n(A)$, respectively. For any $j \in \{1, \cdots, n\}$, let v and w be any unit vectors such that $v \in \operatorname{span}\{u_1, \cdots, u_j\}$ and $w \in \operatorname{span}\{u_j, \cdots, u_n\}$. Then v^*Av and w^*Aw are real numbers satisfying*

$$v^*Av \geqslant \lambda_j(A) \quad \text{and} \quad w^*Aw \leqslant \lambda_j(A).$$

Proof. Because $\{u_1, \cdots, u_n\}$ forms an orthonormal basis for $\mathbb{C}^n$, there exists complex numbers $\alpha_1, \cdots, \alpha_j$ and $\beta_j, \cdots, \beta_n$ satisfying $v = \sum_{i=1}^{j} \alpha_i u_i$ and $w = \sum_{i=j}^{n} \beta_i u_i$. By the same reason, $u_i^* u_j = 0$ whenever $i \neq j$, and $u_i^* u_i = 1$ for all $i \in \{1, \cdots, n\}$. So we have $1 = \|v\|^2 = \sum_{i=1}^{j} |\alpha_i|^2$ and thus

$$\begin{aligned} v^*Av = v^* \sum_{i=1}^{j} A\alpha_i u_i &= v^* \sum_{i=1}^{j} \lambda_i(A)\alpha_i u_i \\ &= \left(\sum_{i=1}^{j} \overline{\alpha_i} u_i^*\right)\left(\sum_{i=1}^{j} \lambda_i(A)\alpha_i u_i\right) \\ &= \sum_{i=1}^{j} \lambda_i(A)\,|\alpha_i|^2\, u_i^* u_i \\ &\geqslant \lambda_j(A) \sum_{i=1}^{j} |\alpha_i|^2 \\ &= \lambda_j(A). \end{aligned}$$

Similarly we have $1 = \|w\|^2 = \sum_{i=j}^{n} |\beta_i|^2$ and thus

$$\begin{aligned} w^*Aw = w^* \sum_{i=j}^{n} A\beta_i u_i &= w^* \sum_{i=j}^{n} \lambda_i(A)\beta_i u_i \\ &= \left(\sum_{i=j}^{n} \overline{\beta_i} u_i^*\right)\left(\sum_{i=j}^{n} \lambda_i(A)\beta_i u_i\right) \\ &= \sum_{i=j}^{n} \lambda_i(A)\,|\beta_i|^2\, u_i^* u_i \\ &\leqslant \lambda_j(A) \sum_{i=j}^{n} |\beta_i|^2 \\ &= \lambda_j(A). \end{aligned}$$

Therefore we have the desired inequalities. □

Claim 2. *Let A be an $n \times n$ Hermitian matrix, and A_k be a $k \times k$ submatrix of A obtained by taking the first k rows and the first k columns. Then $\lambda_j(A_k) \leqslant \lambda_j(A)$ for any $j = 1, \cdots, k$.*

Proof. By assumption there exists a $k \times (n-k)$ matrix B and an $(n-k) \times (n-k)$ matrix C such that $A = \begin{bmatrix} A_k & B \\ B^* & C \end{bmatrix}$. Note that A_k is also Hermitian. Let $v_1, \cdots, v_n$ be eigenvectors of A, each correspoding to eigenvalue $\lambda_1(A), \cdots, \lambda_n(A)$ respectively, and let $w_1, \cdots, w_k$ be eigenvectors of A_k, each correspoding to eigenvalue $\lambda_1(A_k), \cdots, \lambda_k(A_k)$ respectively. Fix any $j \in \{1, 2, \cdots, k\}$ and let V, W be subspaces of $\mathbb{C}^n$ where

$$V = \operatorname{span}\{v_j, \cdots, v_n\}, \quad W = \operatorname{span}\left\{ \begin{bmatrix} w_1 \\ \mathbf{0} \end{bmatrix}, \cdots, \begin{bmatrix} w_j \\ \mathbf{0} \end{bmatrix} \right\}.$$

Then $\dim(V) + \dim(W) = (n - j + 1) + j > n$ so $V \cap W$ must be a nontrivial subspace of $\mathbb{C}^n$. Hence there exists a unit vector $u \in \mathbb{C}^k$ such that $\begin{bmatrix} u \\ \mathbf{0} \end{bmatrix} \in V \cap W$. Now observe that

$$\begin{bmatrix} u \\ \mathbf{0} \end{bmatrix}^* A \begin{bmatrix} u \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} u^* & \mathbf{0} \end{bmatrix} \begin{bmatrix} A_k & B \\ B^* & C \end{bmatrix} \begin{bmatrix} u \\ \mathbf{0} \end{bmatrix} = u^* A_k u.$$

Here as $\begin{bmatrix} u \\ \mathbf{0} \end{bmatrix}$ itself is a unit vector in $\mathbb{C}^n$, and as it is clear that $u \in \operatorname{span}\{w_1, \cdots, w_j\}$, by **Claim 1** we get

$$\lambda_j(A_k) \leqslant u^* A_k u = \begin{bmatrix} u \\ \mathbf{0} \end{bmatrix}^* A \begin{bmatrix} u \\ \mathbf{0} \end{bmatrix} \leqslant \lambda_j(A)$$

which is the inequality desired. □

Claim 3. *Let A be any $n \times n$ Hermitian matrix. Then f can be represented as*

$$f(A) = \sup\left\{ \operatorname{trace}(X^* A X) : X \in \mathbb{C}^{n \times k}, X^* X = I_k \right\}$$

where I_k is the $k \times k$ identity matrix.

Proof. For convenience let S denote the supremum on the right hand side. As A is Hermitian, there exists an $n \times n$ unitary matrix U such that $A = UDU^*$ where $D = \operatorname{diag}\left(\lambda_1(A), \cdots, \lambda_n(A)\right)$. Let $J = \begin{bmatrix} I_k \\ \mathbf{0} \end{bmatrix}$ be the $n \times k$ diagonal matrix with entries of the main diagonal being 1.

Then for $X = UJ$ we have

$$\begin{aligned} \operatorname{trace}(X^* A X) &= \operatorname{trace}(J^* U^* A U J) \\ &= \operatorname{trace}(J^* D J) \\ &= \lambda_1(A) + \cdots + \lambda_k(A) \\ &= f(A) \end{aligned}$$

while $X^* X = J^* U^* U J = J^* J = I_k$. Hence $f(A) \leqslant S$.

On the other hand, let X be any $n \times k$ matrix where $X^* X = I_k$, and let x_i be the i^{th} column vector of X. The fact that $X^* X = I_k$ implies that $\{x_1, \cdots, x_k\}$ is an orthonormal set, so by the Basis Extension Theorem and Gram-Schmidt process we can extend this set into an orthonormal basis of $\mathbb{C}^n$. That is, there exists a unitary matrix Y where if we let y_i to be the i^{th} column vector of Y then $x_1 = y_1, \cdots, x_k = y_k$. Also, as $Y^* Y = I_n$, we have $Y^* X = J$.

Note that $Y^* A Y$ is similar to A, so $\lambda_j(Y^* A Y) = \lambda_j(A)$. Also note that the (i, j)-entry of $Y^* A Y$ is $y_i^* A y_j$, so the $k \times k$ submatrix of $Y^* A Y$ obtained by taking only the first k rows and k columns becomes $X^* A X$. Thus by **Claim 2**, and also by noting that $X^* A X$ is a $k \times k$ matrix, we have

$$\begin{aligned} \operatorname{tr}(X^* A X) &= \lambda_1(X^* A X) + \cdots + \lambda_k(X^* A X) \\ &\leqslant \lambda_1(Y^* A Y) + \cdots + \lambda_k(Y^* A X) \\ &= \lambda_1(A) + \cdots + \lambda_k(A). \end{aligned}$$

Since X was chosen arbitrarily so that $X^* X = I_k$, we have $S \leqslant f(A)$. Therefore as claimed, we have $f(A) = S = \sup\left\{ \operatorname{trace}(X^* A X) : X \in \mathbb{C}^{n \times k}, X^* X = I_k \right\}$. □

Finally now we can show that f is convex. Note that for any given $X \in \mathbb{C}^{n\times k}$, the function defined on the set of $n \times n$ matrices as $A \mapsto \operatorname{trace}(X^*AX)$ is linear. Let A and B be any $n \times n$ Hermitian matrices, and $\lambda \in [0, 1]$. Then by **Claim 3**, we have

$$\begin{aligned} f(\lambda A + (1-\lambda)B) &= \sup\left\{\operatorname{trace}(X^*(\lambda A + (1-\lambda)B)X) : X \in \mathbb{C}^{n\times k}, X^*X = I_k\right\} \\ &= \sup\left\{\lambda \operatorname{trace}(X^*AX) + (1-\lambda)\operatorname{trace}(X^*BX) : X \in \mathbb{C}^{n\times k}, X^*X = I_k\right\} \\ &\leqslant \lambda \sup\left\{\operatorname{trace}(X^*AX) : X \in \mathbb{C}^{n\times k}, X^*X = I_k\right\} \\ &\quad + (1-\lambda)\sup\left\{\operatorname{trace}(X^*BX) : X \in \mathbb{C}^{n\times k}, X^*X = I_k\right\} \\ &= \lambda f(A) + (1-\lambda) f(B). \end{aligned}$$

Therefore by definition, f is convex.